PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-287476
(43)Date of publication of application : 14.10.2004

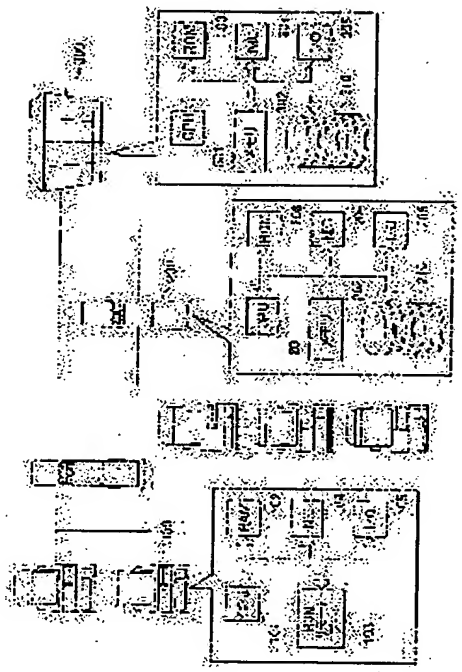
(51)Int.Cl. G06F 12/00

(21)Application number : 2003-075181
(22)Date of filing : 19.03.2003

(71)Applicant : HITACHI LTD
(72)Inventor : KIMURA SHINJI
OSHIMA SATOSHI
HASHIMOTO TAKASHI

(54) CACHE CONTROL FOR NODE DEVICE

(57)Abstract:
PROBLEM TO BE SOLVED: To improve security of data against the theft of a node device and unauthorized access in a computer system comprising a client, the node device and a storage device.
SOLUTION: In the computer system capable of transferring data through the node device between the client and storage device, cache in the node device is controlled in the following method. In the first control, the attribute of cache propriety is given on data stored in the storage device, and the node device carries out relay to the client without performing cache on cache-disabled data. In the second control, the node device enciphers data in caching to a disk. In the third control, the cache-disabled data are directly transmitted and received without passing through the node device. With this control, cache to the node device can be limited to ensure security.



LEGAL STATUS

[Date of request for examination] 04.10.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2004-287476
(P2004-287476A)
(43)公開日　平成16年10月14日(2004.10.14)

(51)Int.Cl.[7]	F I		テーマコード（参考）
G06F 12/00	G06F 12/00	546K	5B082
	G06F 12/00	537Z	

審査請求 未請求　請求項の数 28　OL　（全 25 頁）

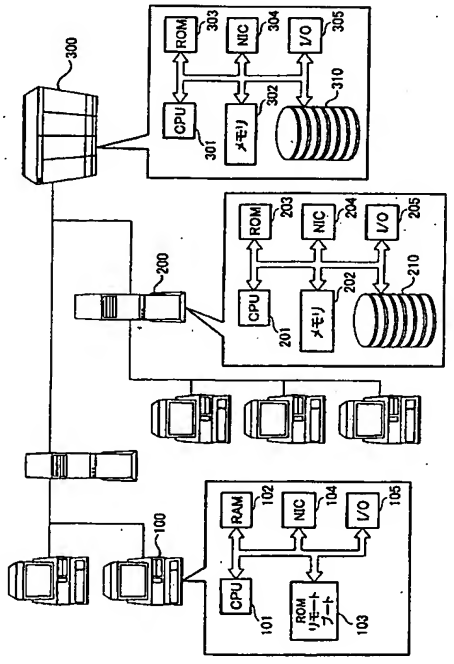
(21)出願番号　特願2003-75181 (P2003-75181)
(22)出願日　平成15年3月19日 (2003.3.19)

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(74)代理人　110000028
特許業務法人明成国際特許事務所
(72)発明者　木村　信二
神奈川県川崎市麻生区王禅寺１０９９番地
株式会社日立製作所システム開発研究所内
(72)発明者　大島　劼
神奈川県川崎市麻生区王禅寺１０９９番地
株式会社日立製作所システム開発研究所内

最終頁に続く

(54)【発明の名称】ノード装置におけるキャッシュ制御

(57)【要約】
【課題】クライアント、ノード装置、ストレージ装置からなる計算機システムにおいて、ノード装置の盗難、不正アクセスに対するデータのセキュリティを向上する。
【解決手段】クライアントおよびストレージ装置間で、ノード装置を介してデータを授受可能な計算機システムにおいて、ノード装置におけるキャッシュを次の方法で制御する。第１の制御では、ストレージ装置に格納されているデータについてキャッシュ可否の属性を付与しておき、ノード装置は、キャッシュ不可とされているデータについては、キャッシュを行わずにクライアントへの中継を行う。第２の制御では、ノード装置は、ディスクへのキャッシュ時にデータの暗号化を行う。第３の制御では、キャッシュ不可のデータについては、ノード装置を介さずに直接送受信する。これらの制御により、ノード装置へのキャッシュを制限でき、セキュリティを確保することができる。
【選択図】　図１



【特許請求の範囲】
【請求項１】
ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、
前記ストレージ装置またはクライアントにおいて、
前記ノード装置によって中継されるデータについて、前記ディスク装置へのキャッシュの可否を表す属性情報を前記ノード装置に対して出力するステップを実行し、
前記ノード装置において、
前記属性情報に基づいて、前記中継するデータの前記ディスク装置へのキャッシュ可否を判断するステップと、
キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継するステップとを実行するキャッシュ制御方法。
【請求項２】
ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、
前記ノード装置において、
中継すべきデータを、前記ストレージ装置またはクライアントから入力するステップと、
前記中継するデータを暗号化して前記ディスク装置に書き込むステップと、
前記ディスク装置に保持されているデータの読出コマンドに応じて、該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力するステップとを実行するキャッシュ制御方法。
【請求項３】
ストレージ装置、キャッシュ用のディスク装置を備えるノード装置、およびクライアントが接続された計算機システムにおいて、前記ディスク装置へのキャッシュを制御するキャッシュ制御方法であって、
前記ノード装置において、
前記ストレージ装置またはクライアント装置から受信したデータを、所定のタイミングで前記ディスク装置にキャッシュするステップを実行し、
前記ストレージ装置またはクライアントの一方において、
提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断するステップと、
キャッシュ可と判断されたデータについては、前記ノード装置に送信するステップと、
キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記ストレージ装置またはクライアントの他方に直接データを送信するステップとを実行するキャッシュ制御方法。
【請求項４】
キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置であって、
前記中継するデータについて、前記ディスク装置へのキャッシュの可否を表す属性情報を入力する属性情報入力部と、
該属性情報に基づいて、前記中継するデータの前記ディスク装置へのキャッシュ可否を判断する判断部と、
キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継するキャッシュ制御部とを備えるノード装置。
【請求項５】
請求項４記載のノード装置であって、
前記ディスク装置の他に、キャッシュ用の揮発性メモリを有し、
前記キャッシュ制御部は、

前記キャッシュ可否に関わらず前記データを前記揮発性メモリにキャッシュするメモリ制御部と、
前記揮発性メモリから前記ディスク装置への書き込みを行うための所定の条件が成立した場合に、前記キャッシュ否と判断されたデータを除き、該揮発性メモリに保持されたデータを前記ディスク装置に移転する移転制御部とを備えるノード装置。
【請求項６】
キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置であって、
前記中継するデータを暗号化して前記ディスク装置に書き込む暗号化部と、
前記ディスク装置に保持されているデータの読出コマンドに応じて、該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力する復号部とを備えるノード装置。
【請求項７】
請求項６記載のノード装置であって、
前記中継するデータについて、暗号化の要否を表す属性情報を入力する属性情報入力部と、
該属性情報に基づいて、前記中継するデータの暗号化の要否を判断する判断部と、
暗号化要と判断されたデータについては、前記暗号化部および復号部を用いて前記ディスク装置へのキャッシュを行うキャッシュ制御部とを備えるノード装置。
【請求項８】
請求項６記載のノード装置であって、
前記ストレージ装置から前記暗号化および復号に使用する鍵データを受信し、揮発性メモリに管理する鍵データ管理部を備えるノード装置。
【請求項９】
請求項４～８いずれか記載のノード装置であって、
前記属性情報は、前記データの中継時に、該データと関連付けて入力されるノード装置。
【請求項１０】
請求項４～８いずれか記載のノード装置であって、
前記属性情報入力部は、前記ストレージ装置におけるデータの格納ブロックと前記属性情報との関係を予め取得し、管理するノード装置。
【請求項１１】
請求項４～８いずれか記載のノード装置であって、
前記属性情報は、前記クライアントにおける異常発生を通知する情報であるノード装置。
【請求項１２】
ディスク装置を有するノード装置を介して、クライアントに、データを提供するストレージ装置であって、
前記データについて、前記ディスク装置へのキャッシュの可否、または前記ディスク装置への書き込み時の暗号化の要否を表す属性情報を管理する属性情報管理部と、
前記ノード装置に、該属性情報を通知する属性情報通知部とを備えるストレージ装置。
【請求項１３】
請求項１２記載のストレージ装置であって、
前記属性情報通知部は、前記データの提供時に、該データと関連づけて前記属性情報を通知するストレージ装置。
【請求項１４】
請求項１２記載のストレージ装置であって、
前記属性情報通知部は、前記データの格納ブロックと前記属性情報との関係をデータの提供に先立って前記ノード装置に通知するストレージ装置。
【請求項１５】
請求項１２記載のストレージ装置であって、
前記ディスクへの書き込み時にデータを暗号化可能な複数のノード装置と接続されており

いずれかの前記ノード装置に対して前記暗号化に用いる鍵データの出力を行うべき所定の条件が満たされたか否かを判断する条件判断部と、
該ノード装置に対して前記鍵データを送信する鍵データ送信部とを備えるストレージ装置。

【請求項１６】
ディスク装置を有するノード装置を介してストレージ装置にデータを書き込み可能なクライアントであって、
前記ディスク装置へのキャッシュの禁止、または前記ディスク装置への書き込み時の暗号化を前記ノード装置に要求するための所定の条件が、満たされたか否かを判断する条件判断部と、
該条件が満たされた時に、前記ノード装置に対して、前記要求を表す属性情報を出力する禁止通知部と、
該属性情報の出力と併せて、または該属性情報の出力後に、前記データをノード装置に送信するデータ送信部とを備えるクライアント。

【請求項１７】
請求項１６記載のクライアントであって、
前記所定の条件は、該クライアントにおけるソフトウェアの動作異常時であるクライアント。

【請求項１８】
キャッシュ用のディスク装置を備えるノード装置を介して、他の計算機にデータを提供するデータ提供装置であって、
前記提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断する判断部と、
キャッシュ可と判断されたデータについては前記ノード装置に送信し、キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記他の計算機に送信する送信制御部とを備えるデータ提供装置。

【請求項１９】
請求項１８記載のデータ提供装置であって、
前記提供するデータについて、キャッシュの可否を表す属性情報を予め管理する属性情報管理部を備え、
前記判断部は、該属性情報に基づいてキャッシュの可否を判断するデータ提供装置。

【請求項２０】
請求項１８記載のデータ提供装置であって、
前記判断部は、該データ提供装置におけるソフトウェアの動作異常時に、キャッシュ否と判断するデータ提供装置。

【請求項２１】
請求項１８記載のデータ提供装置であって、
前記他の計算機にデータを送信するためのアドレス情報を管理するアドレス管理部を備え、
前記送信制御部は、前記キャッシュ否と判断された場合には、データの送信先を前記ノード装置のアドレスから前記アドレス管理部で管理されているアドレス情報に切り換えるデータ提供装置。

【請求項２２】
ストレージ装置と複数のクライアントとの間で、データの授受を中継するノード装置であって、
いずれかの前記クライアントから前記ストレージ装置内のデータの読出コマンドを受け付ける要求受付部と、
該読出コマンドを発信した前記クライアントを特定可能な情報を、該読出コマンドとともに、前記ストレージ装置に送信する要求送信部とを備えるノード装置。

【請求項２３】
キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置の動作を制御するコンピュータプログラムであって、
前記中継するデータについて、キャッシュの可否を表す属性情報を入力する機能と、
該属性情報に基づいて、前記中継するデータのキャッシュ可否を判断する機能と、
キャッシュ否と判断されたデータについては、前記ディスク装置へのキャッシュを行わずに中継する機能とを該ノード装置において実現するためのコンピュータプログラム。
【請求項２４】
キャッシュ用のディスク装置を備え、ストレージ装置とクライアントとの間で、データの授受を中継するノード装置の動作を制御するためのコンピュータプログラムであって、
前記中継するデータを暗号化して前記ディスク装置に書き込む機能と、
該ディスク装置から読み出したデータを復号して前記ストレージ装置またはクライアントに出力する機能とを該ノード装置において実現するためのコンピュータプログラム。
【請求項２５】
ディスク装置を有するノード装置を介して、クライアントに、データを提供するストレージ装置の動作を制御するためのコンピュータプログラムであって、
前記データについて、前記ディスク装置へのキャッシュの可否、または前記ディスク装置への書き込み時の暗号化の要否を表す属性情報を管理する機能と、
前記ノード装置に、該属性情報を通知する機能とを該ストレージ装置において実現するためのコンピュータプログラム。
【請求項２６】
ディスク装置を有するノード装置を介してストレージ装置にデータを書き込み可能なクライアントの動作を制御するためのコンピュータプログラムであって、
前記ディスク装置へのキャッシュの禁止、または前記ディスク装置への書き込み時の暗号化を、前記ノード装置に要求するための所定の条件が満たされたか否かを判断する機能と、
該条件が満たされた時に、前記ノード装置に対して、前記要求を表す属性情報を出力する機能と、
該属性情報の出力と併せて、または該属性情報の出力後に、前記データをノード装置に送信する機能とを該クライアントにおいて実現するためのコンピュータプログラム。
【請求項２７】
キャッシュ用のディスク装置を備えるノード装置を介して、他の計算機にデータを提供するデータ提供装置の動作を制御するコンピュータプログラムであって、
前記提供するデータについて、所定の条件に基づいて、前記ディスク装置へのキャッシュの可否を判断する機能と、
キャッシュ否と判断されたデータについては、前記ノード装置を介さずに前記他の計算機に直接データを送信する機能とを該データ提供装置において実現するためのコンピュータプログラム。
【請求項２８】
ストレージ装置と複数のクライアントとの間で、データの授受を中継するノード装置の動作を制御するコンピュータプログラムであって、
いずれかの前記クライアントから前記ストレージ装置内のデータの読出コマンドを受け付ける機能と、
該読出コマンドを発信した前記クライアントを特定可能な情報を、該読出コマンドとともに、前記ストレージ装置に送信する機能とを該ノード装置において実現するためのコンピュータプログラム。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、ストレージ装置、クライアント間でデータを中継するノード装置におけるキャ

ッシュの制御に関する。

【０００２】

【従来の技術】

企業などでは、複数のクライアントコンピュータ（以下、単にクライアントと称する）をネットワークで接続した大規模な計算機システムが構築されている。このような計算機システムでは、各クライアントで業務を遂行する際に参照する種々のデータやプログラムを供給するため、ストレージ装置と呼ばれる大容量の記憶装置が設けられている。ストレージ装置は、大容量のハードディスクを備えており、各クライアントのユーザは、自己に割り当てられた領域、または共有の領域に対して、データやプログラムなどの格納・読み出しを行う。通常、ストレージ装置へのアクセスには、ＳＣＳＩ（Ｓｍａｌｌ Ｃｏｍｐｕｔｅｒ Ｓｙｓｔｅｍ Ｉｎｔｅｒｆａｃｅ）と呼ばれるプロトコルが使用される。近年では、ＳＣＳＩ用のコマンドを、いわゆるＩＰパケットに分割してネットワーク上を転送するためのｉＳＣＳＩと呼ばれるプロトコルが提案されている。ｉＳＣＳＩを用いることにより、従来よりも更に多数のクライアント、および遠隔地のクライアントがストレージ装置にアクセス可能な環境が提供されつつある。

【０００３】

ストレージ装置の活用方法の一つとして、リモートブートと呼ばれるクライアントの起動方法が提案されている。リモートブートでは、起動に必要なオペレーティングシステム、環境設定ファイルなどの各種ファイルが、予めストレージ装置の各ユーザ用の領域に格納されている。各クライアントは、ネットワークを介してこれらのファイルをストレージ装置から読み込むことにより、起動することができる。

【０００４】

リモートブートする場合、クライアントは、ストレージ装置へのアクセスを実現するための比較的小さなブートプログラムをＲＯＭに記憶していれば十分であり、ハードディスクを備えている必要がない。ハードディスクに格納されるべきデータはストレージ装置で一元管理されるから、仮にクライアントコンピュータが破損、盗難などに遭っても、重要なデータの遺失、漏洩を逃れることができるというセキュリティ上の利点がある。また、各ユーザは、ネットワーク上に接続されたいずれのクライアントも、一定の環境設定で起動させることができるという実用上の利点もある。リモートブートについては、例えば、特許文献１～３に開示されている。

【０００５】

ストレージ装置を用いる場合、多数のクライアントによるストレージ装置へのアクセスの集中を緩和するため、クライアントとストレージ装置との間に、キャッシュ用のメモリを備えたノード装置を設けることがある。かかるシステムでは、ノード装置にキャッシュ済みのデータは、ストレージ装置にアクセスするまでなくクライアントに供給可能であるため、ストレージ装置へのアクセスが緩和される。

【０００６】

【特許文献１】

特開平６－３３２７１６号公報

【特許文献２】

特開２００１－７５８５３号公報

【特許文献３】

特開２０００－２５９５８３号公報

【０００７】

【発明が解決しようとする課題】

しかし、ノード装置を用いたシステムでは、このノード装置がセキュリティ上、脆弱になり易いという課題があった。一般に、ストレージ装置は、企業の機密データを一元的に保持しているため、非常に厳しくセキュリティ管理されている。例えば、ストレージ装置が設置された部屋への立ち入りは厳しく制限されており、ソフトウェア的にもファイアウォールなどでストレージ装置へのアクセスは厳しく制限されている。一方、クライアントは

、ハードディスクを備えないディスクレスコンピュータとすることにより、データの漏洩を防止することができる。仮に、クライアントがハードディスクを備えている場合でも、各ユーザは、自己が使用する装置については盗難や不正使用に遭わないよう気を配るため、一定レベルでセキュリティは確保される。

【０００８】

これに対し、ノード装置に対しては、ストレージ装置ほどセキュリティが確保された環境が確保されておらず、各ユーザもセキュリティに気を配らないことが多い。かかる環境下で、ノード装置には、ストレージ装置とクライアントの間で授受される機密データがキャッシュされている。従って、ノード装置の盗難や不正アクセスによって、機密データが漏洩する可能性があった。かかる機密データの一つとして、例えば、動作に支障が生じた時にクライアントが出力するコアダンプが挙げられる。コアダンプは、通常のコンピュータでは、内蔵のハードディスクに書き出されるが、ディスクレスコンピュータではノード装置にキャッシュされた後、ストレージ装置に書き出される。コアダンプには、ストレージ装置にアクセスするためのパスワードなど、コンピュータに保存されていた種々の機密情報が含まれており、これらの情報の漏洩は、ストレージ装置への不正アクセスを許容する原因となり得る。本発明は、これらの課題に鑑み、ノード装置におけるセキュリティの向上を図ることを目的とする。

【０００９】

【課題を解決するための手段】

本発明では、ストレージ装置、ノード装置、およびクライアントが接続された計算機システムにおいて、ノード装置に備えられたキャッシュ用のディスク装置へのキャッシュを次の方法で制御する。この制御は、ノード装置の揮発性メモリへのキャッシュにも同様に適用してもよいし、ディスク装置へのキャッシュ時にのみ適用してもよい。

【００１０】

＜第１の制御方法＞　第１の制御方法では、ディスク装置へのキャッシュの可否を表す属性情報を用いる。ストレージ装置またはクライアントは、ノード装置によって中継されるデータについて、上述の属性情報をノード装置に対して出力する。ノード装置は、この属性情報に基づいて、キャッシュ否と判断されたデータについては、ディスク装置へのキャッシュを行わずに中継する。こうすることにより、機密データについてはノード装置でのキャッシュを禁止することができ、セキュリティを確保することができる。

【００１１】

ノード装置がディスク装置の他に、キャッシュ用の揮発性メモリを有している場合、キャッシュ否のデータについては揮発性メモリへのキャッシュを禁止してもよい。また、揮発性メモリに対しては、キャッシュ可否に関わらずキャッシュを行い、揮発性メモリからディスク装置への書き込みは、キャッシュ否と判断されたデータを除いて行うようにしてもよい。

【００１２】

キャッシュ可否を制御するための属性情報は、種々のタイミングで取得することができる。例えば、データの中継時に、このデータと関連付けて属性情報を入力するようにしてもよい。ヘッダ等に属性情報を含めることでデータと属性情報を一体として入力してもよいし、中継されるデータと分けて属性情報のみを入力してもよい。　かかる態様は、例えば、ストレージ装置において、データごとに属性情報を予め管理しておき、ノード装置に、この属性情報を通知する機能をもたせることにより、実現することができる。

【００１３】

ストレージ装置において、データの格納ブロックと対応づけて属性情報が管理されている場合、ノード装置は、この対応を予め取得し、管理しておいてもよい。こうすれば、データの中継時にストレージ装置から属性情報を受け取るまでなく、キャッシュの可否を判断することができる。

【００１４】

属性情報は、必ずしもデータ単位でキャッシュの可否を通知する情報である必要はない。

例えば、クライアントにおけるソフトウェアの動作異常発生を通知する情報を属性情報として用いても良い。ノード装置は、この通知を受信した後は、そのクライアントからのデータのキャッシュを行わないというように、キャッシュ可否の動作を切り替えることができる。同様に、クライアントの動作が正常に復帰したことを通知する情報を属性情報として用いても良い。このように、属性情報は、ノード装置におけるキャッシュ可否を切り替えるタイミングを通知する情報とすることもできる。

【００１５】

かかる通知は、キャッシュの禁止をノード装置に要求するための所定の条件が満たされた時に、ノード装置に対して属性情報を出力する機能をクライアントに持たせることにより実現できる。クライアントは、この属性情報の出力と併せて、または該属性情報の出力後に、データをノード装置に送信するようにすればよい。キャッシュの禁止を要求するための条件は、例えば、上述の動作異常の他、予め設定された特定のアプリケーションの起動時など種々の設定が可能である。

【００１６】

＜第２の制御方法＞　本発明の第２の制御方法では、ノード装置において、中継すべきデータを暗号化してディスク装置に書き込むことでキャッシュする。キャッシュされたデータの読出コマンドを受けた時には、ディスク装置から読み出したデータを復号してストレージ装置またはクライアントに出力する。暗号化は、ディスク装置への書き出し時に完了していればよく、例えば、書き出し直前に行ってもよいし、データの受信時に行ってもよい。こうすることで、ディスク装置が盗難に遭っても、データのセキュリティを確保することができる。

【００１７】

第２の制御方法においては、必ずしも全てのデータを暗号化する必要はない。例えば、中継するデータについて、暗号化の要否を表す属性情報を入力し、この属性情報に基づいて、暗号化の要否を切り替えても良い。属性情報は、第１の制御方法でキャッシュの可否を制御するために例示した種々の情報を適用することができる。こうすれば、不要な暗号化を抑制することができ、ノード装置での処理負担を軽減することができる。第１の制御方法と同様、属性情報は、暗号化の要否を切り替えるタイミングを通知する情報としてもよい。

【００１８】

第２の制御方法においては、暗号化および復号に用いる鍵データを揮発性メモリに管理することが好ましい。こうすれば、ノード装置が盗難等にあった場合には、鍵データが消失するため、ディスク装置に記憶されたデータの復号が不能となる。この鍵データは、例えば、ストレージ装置がノード装置に与えるようにしてもよいし、ノード装置自身が生成するようにしてもよい。ストレージ装置から鍵データを供給する場合、ストレージ装置は、各ノード装置に与えられる鍵データに重複が生じないよう管理することが好ましい。こうすることにより、与えられた鍵データは他のノード装置には適用できなくなるため、ノード装置へのなりすましによって鍵データが取得された場合でも、セキュリティを確保することが可能となる。ストレージ装置からノード装置への鍵データの供給は、例えば、ノード装置が新たに稼働し始めた場合や、ノード装置の管理者がストレージ装置に対して鍵データの発行を要求した場合など、種々の条件下で行うことができる。

【００１９】

＜第３の制御方法＞　第３の制御方法は、ストレージ装置およびクライアントの一方から他方へのデータ送信時において、データの発信側となるデータ提供装置が、データの送信先を切り換えることにより実現される。データ提供装置は、提供するデータについて、ディスク装置へのキャッシュの可否を判断する。キャッシュ可否の判断条件は、例えば、予めデータごとに設定しておいてもよいし、「ソフトウェアの動作異常時」など特定の事象やタイミングとの関係で設定してもよい。データ提供装置はキャッシュ可と判断されたデータについては、ノード装置に送信し、ノード装置は、このデータをディスク装置にキャッシュする。キャッシュ否と判断されたデータについては、ノード装置を介さずに、受信

側の装置に、直接データを送信する。こうすることにより、ノード装置では特別な制御を要することなく、キャッシュを回避することができる。

【００２０】

第３の制御方法を実現するため、データ提供装置は、ノード装置とは別に、受信側となる計算機のアドレス情報を予め管理しておいてもよい。データ提供装置は、キャッシュ否と判断された場合には、データの送信先をノード装置のアドレスから、予め管理されているアドレス情報に切り換えることにより、ノード装置を迂回した通信を実現することができる。この態様は、受信側となる計算機のアドレス情報が予め固定されている場合、例えば、クライアントからストレージ装置へのデータ送信時に有用である。

【００２１】

別の態様として、ノード装置は、ストレージ装置内のデータの読出コマンドを受け付けた時、この読出コマンドを発信したクライアントを特定可能な情報、例えば、アドレス情報、クライアント名、クライアントの識別情報などを、ストレージ装置に送信するようにしてもよい。こうすることでストレージ装置は、ノード装置を迂回する場合のデータの送信先を特定することができる。ストレージ装置が、ノード装置を迂回してクライアントにデータを送信する場合、ノード装置に対して、データの送信が完了した通知を行うようにしてもよい。この通知がノード装置からストレージ装置に出力した読出コマンドに対する応答となり、ノード装置を応答待ちの状態から解放することができる。

【００２２】

本発明は、上述したノード装置、ストレージ装置、クライアントに限らず種々の態様で構成することが可能である。例えば、これらを接続した計算機システムとして構成してもよい。また、ノード装置におけるキャッシュを制御する制御方法として構成してもよい。コンピュータによりこの制御を実現するためのコンピュータプログラムおよびこれを記録した記憶媒体として構成してもよい。ここで、記憶媒体としては、フレキシブルディスクやＣＤ－ＲＯＭ、ＤＶＤ、光磁気ディスク、ＩＣカード、ＲＯＭカートリッジ、パンチカード、バーコードなどの符号が印刷された印刷物、コンピュータの内部記憶装置（ＲＡＭやＲＯＭなどのメモリ）および外部記憶装置などコンピュータが読取り可能な種々の媒体を利用できる。

【００２３】

【発明の実施の形態】

本発明の実施の形態について以下の順序で説明する。

Ａ．第１実施例；
Ａ１．システム構成；
Ａ２．機能ブロック；
Ａ３．テーブルのデータ構造；
Ａ４．データ読出処理；
Ａ５．データ書込処理；
Ａ６．変形例；
Ｂ．第２実施例；
Ｂ１．機能ブロック；
Ｂ２．データ読出処理；
Ｂ３．データ書込処理；
Ｃ．第３実施例；
Ｃ１．機能ブロック；
Ｃ２．データ読出処理；
Ｃ３．データ書込処理；

【００２４】

Ａ．第１実施例；
Ａ１．システム構成；

図１は第１実施例としての計算機システムの構成を示す説明図である。この計算機システ

ムでは、ストレージ装置３００に、複数のノード装置２００および計算機１００がＩＰパケットを授受可能なネットワークで接続されている。ネットワークとしては、ＬＡＮ（Ｌｏｃａｌ Ａｒｅａ Ｎｅｔｗｏｒｋ）、イントラネット、インターネットなどを適用することができる。

【００２５】

ストレージ装置３００は、図示するハードウェア構成を有するコンピュータとして構成されており、大容量のディスク装置３１０に蓄積されたデータをノード装置２００経由で各計算機１００に提供することができる。また、各計算機１００はハードディスク３１０にデータを格納することもできる。ＮＩＣ（Ｎｅｔｗｏｒｋ Ｉｎｔｅｒｆａｃｅ Ｃａｒｄ）３０４は、ネットワークを介してデータを授受するためのインタフェースである。Ｉ／Ｏ３０５は、入出力装置用のインタフェースである。

【００２６】

ストレージ装置３００では、ＣＰＵ３０１が、ディスク装置３１０およびＲＯＭ３０３に記憶された制御用のプログラムに従って、データの授受を制御する。メモリ３０２は、このプログラムの動作時に使用される主記憶である。

【００２７】

ノード装置２００は、図示するハードウェア構成を有するコンピュータとして構成されており、ストレージ装置３００と計算機１００の間で授受されるデータを中継する機能を奏する。ディスク装置２１０およびメモリ２０２は、中継されるデータを一時的に蓄積するキャッシュのために利用される。ＮＩＣ（Ｎｅｔｗｏｒｋ Ｉｎｔｅｒｆａｃｅ Ｃａｒｄ）２０４は、ネットワークを介してデータを授受するためのインタフェースである。Ｉ／Ｏ２０５は、入出力装置用のインタフェースである。

【００２８】

ノード装置２００では、ＣＰＵ２０１が、ディスク装置２１０およびＲＯＭ２０３に記憶された制御用のプログラムに従って、データの中継およびキャッシュを制御する。メモリ２０２は、キャッシュの他、このプログラムの動作時にも使用される。本実施例では、中継するデータによってキャッシュの可否が切り替えられる。キャッシュが禁止されているデータについては、ノード装置２００は、キャッシュを行わずに中継する。キャッシュの禁止については、二通りの制御が適用可能である。第１の態様は、キャッシュが禁止されているデータに関しては、ノード装置２００は一切、キャッシュを行わないという制御方法である。第２の態様は、ノード装置２００は揮発性のメモリ２０２には、キャッシュ可否の属性に関わらずキャッシュを行い、メモリ２０２からディスク装置２１０にデータを移転する際に、キャッシュの可否を判断する制御方法である。本実施例では、第１の態様でキャッシュの可否を制御するものとして説明し、後で変形例として第２の態様について説明する。

【００２９】

計算機１００は、ハードディスクを備えないディスクレスコンピュータである。内部には、ＣＰＵ１０１、ＲＡＭ１０２、ＲＯＭ１０３、ＮＩＣ１０４およびＩ／Ｏ１０５が備えられている。計算機１００は、起動時には、必要なオペレーティングシステムなどのファイルをストレージ装置３００から読み込むことで、リモートブートする。ＲＯＭ１０３には、リモートブート用のプログラムが予め記憶されている。

【００３０】

ストレージ装置３００、ノード装置２００、計算機１００の動作を制御するためのプログラムは、ＣＤ－ＲＯＭなどの記録媒体によってストレージ装置に供給してもよい。

【００３１】

Ａ２．機能ブロック；

図２は計算機システムを構成する各装置内の機能ブロックを示す説明図である。本実施例では、図示する各機能ブロックは、それぞれコンピュータプログラムをインストールすることによりソフトウェア的に構成される。各機能ブロックは、ハードウェア的に構成することも可能である。

【００３２】
計算機１００の機能ブロックは、次の通りである。ＴＣＰ／ＩＰモジュール１２０は、ネットワークを介してＩＰパケットを通信する。ＳＣＳＩモジュール１２２は、ストレージ装置３００に対してデータの読み出し、書き込み用のコマンド送受信などを行う。ｉＳＣＳＩモジュール１２１は、ＳＣＳＩコマンドとＩＰパケットの変換を行うことによって、ＳＣＳＩコマンドをＴＣＰ／ＩＰで送受信可能とする。ブートプログラム１３０は、計算機１００でリモートブートを実現する。ブートプログラム１３０は、計算機１００の初期化を行う機能、起動に必要なオペレーティングシステムなどのファイルをストレージ装置から読み込む機能、これらのファイルを起動してブートする機能などを実現することができる。ＩＤ入力モジュール１３１は、リモートブートを行う際に、ストレージ装置３００に送信すべきユーザＩＤ、パスワードを入力する。本実施例では、これらのモジュールは、ＲＯＭ１０３に記憶されているプログラムによって実現される。

【００３３】
上述の各機能ブロックによってリモートブートが完了すると、計算機１００には、ストレージ装置３００から供給されたファイルによってオペレーティングシステム１４０が起動する。オペレーティングシステム１４０には、計算機１００の動作に異常が生じた時に、種々の情報をコアダンプとしてストレージ装置３００に書き出す機能を奏する異常制御モジュール１４１が備えられている。計算機１００上の各処理に利用されるアプリケーション１５０は、ストレージ装置３００から供給され、オペレーティングシステム１４０の上で稼働する。

【００３４】
本実施例では、クライアントごとに仮想ボリュームが与えられる。仮想ボリュームとは、ストレージ装置３００のディスク装置３１０内に設けられた物理的な領域ではなく、各クライアントのデータを管理するための論理的なディスクを意味する。仮想ボリュームは、ユーザＩＤに対応づけて設定されており、ユーザが、計算機１００でユーザＩＤを入力することによって、計算機１００と対応づけられる。以下、本実施例では、「クライアント」という用語は、単にハードウェアとしての計算機ではなく、ユーザがログインし、仮想ボリュームと対応づけられた状態にある計算機を意味するものとする。

【００３５】
ノード装置２００には、ネットワークを介してＳＣＳＩコマンドを授受するため、ＴＣＰ／ＩＰモジュール２２０、ｉＳＣＳＩモジュール２２１、ＳＣＳＩモジュール２２２が備えられている。ノード装置２００には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってキャッシュ制御モジュール２３０が実現されている。

【００３６】
キャッシュ制御モジュール２３０は、キャッシュ属性管理部２３４およびキャッシュ管理テーブル２３２を参照して、ノード装置２００におけるキャッシュを制御する。キャッシュ管理テーブル２３２とは、クライアントから指定された仮想ブロックと、ノード装置２００におけるデータの格納場所、即ちキャッシュブロックとを対応づけるテーブルである。キャッシュ属性管理部２３４とは、仮想ブロックの各データについて、キャッシュの可否を示す属性情報を管理するテーブルである。これらのテーブルのデータ構造については後述する。

【００３７】
ストレージ装置３００には、ネットワークを介してＳＣＳＩコマンドを授受するため、ＴＣＰ／ＩＰモジュール３２０、ｉＳＣＳＩモジュール３２１、ＳＣＳＩモジュール３２２が備えられている。ストレージ装置３００には、所定のオペレーティングシステムがインストールされており、このオペレーティングシステム上で稼働するアプリケーションによってアクセス制御モジュール３３０が実現されている。

【００３８】
アクセス制御モジュール３３０は、ユーザ管理テーブル３３４、物理ブロック管理テーブ

ル３３２を参照して、ディスク装置３１０のデータの読み書きを制御する。ユーザ管理テーブル３３４は、ユーザＩＤと仮想ボリュームとを対応づけるテーブルである。物理ブロック管理テーブル３３２とは、仮想ブロックと物理ブロックとを対応づけるテーブルである。アクセス制御モジュール３３０は、計算機１００の起動時に、計算機１００からユーザＩＤを取得する。このユーザＩＤに基づいてユーザ管理テーブル３３４を参照することで、ユーザに対応する仮想ボリュームを計算機１００に割り当てる。アクセス制御モジュール３３０は、以後、クライアントから指定された仮想ブロックに対して、物理ブロック管理テーブル３３２を参照して、ディスク装置３１０のデータの読み書きを制御することができる。

【００３９】

キャッシュ管理テーブル３３３は、仮想ブロックごとにキャッシュの可否を記録したテーブルである。アクセス制御モジュール３３０は、キャッシュ属性管理テーブル３３３の内容をノード装置２００に通知することで、ノード装置２００におけるキャッシュの可否を制御する。

【００４０】

Ａ３．テーブルのデータ構造；

図３はユーザ管理テーブル３３４の構造を示す説明図である。ユーザ管理テーブル３３４は、ユーザＩＤ、パスワード、仮想ボリューム、属性、ＩＰアドレスを対応づけて管理する。属性とは、仮想ボリュームに対する他のユーザのアクセスの可否を表している。「専有」は他のユーザのアクセスが禁止されていることを意味し、「共有」は許可されていることを意味する。ＩＰアドレスは、ユーザがログインした計算機によって変動する。属性、ＩＰアドレスは、ユーザ管理テーブル３３４から省略してもよい。

【００４１】

図４は物理ブロック管理テーブル３３２の構造を示す説明図である。図の上方にデータ構造を示し、下方に仮想ブロックと物理ブロックとの対応関係を模式的に示した。物理ブロック管理テーブル３３２は、仮想ブロックと物理ブロックの対応関係を示すテーブルであり、仮想ボリューム単位で設けられている。図中には、仮想ボリュームＶＤａに対応するテーブルを例示した。このテーブルでは、仮想ボリュームＶＤａの仮想ブロックＢＬ０～ＢＬｎに対して、実際にデータが格納されている物理ブロックが記憶されている。物理ブロックは、物理ボリュームとブロック番号との組み合わせで定義される。

【００４２】

図に例示したテーブルは、下方に模式的に示した対応関係を表している。つまり、仮想ブロックＢＬ０のデータは、物理ボリュームＰＤ１１のブロックＢＬ０に格納されている。同様にして、仮想ブロックＢＬ１～ＢＬｎの各データは、それぞれ物理ボリュームＰＤ１１、ＰＤ１２のいずれかのブロックに格納されている。図では、ディスク装置３１０が、パーティションによって複数の物理ボリュームに分割されている場合を例示したが、単一の物理ボリュームとして構成されていてもよい。この場合、物理ブロック管理テーブル３３２からは、物理ボリュームに関するデータを省略しても構わない。

【００４３】

図５はキャッシュ管理テーブル２３２の構造を示す説明図である。このテーブルは、仮想ブロックと、ノード装置２００におけるキャッシュブロックとを対応づけている。ディスク書込の欄については、後述する。ノード装置２００におけるキャッシュの格納場所には、揮発性のメモリ２０２と、ディスク装置２１０の２種類が存在する。従って、本実施例では、格納場所とブロック番号で、キャッシュブロックを特定するものとした。

【００４４】

図６はキャッシュ属性管理テーブル３３３の構造を示す説明図である。キャッシュ属性管理テーブル３３３は、データの読み出しと書き込みに分けて用意されており、仮想ボリュームごとにキャッシュの可否を表す情報を格納している。本実施例では、ディスク装置が複数の物理ボリュームに分割されているため、物理ボリューム単位でキャッシュの可否を設定可能とした。例えば、図の例では、仮想ボリュームＶＤａの読み出しについて、物理

ボリュームＰＤ１１に格納されたデータについてはキャッシュ可、物理ボリュームＰＤ１２に格納されたデータについてはキャッシュ不可と設定されている。

【００４５】

図の例において、物理ボリュームＰＤ１２は、仮想ボリュームＶＤａ、ＶＤｂに共有されている。本実施例では、このように複数の仮想ボリュームによる物理ボリュームの共有を許容してもよい。この場合、共有される物理ボリュームに格納されたデータは、各仮想ボリュームに対応した複数のクライアントで共有されることになる。図中の例では、共有されている物理ボリュームＰＤ１２について、キャッシュ属性は不可と設定されている場合を例示した。これに対し、仮想ボリュームＶＤａ、ＶＤｂで、共有されている物理ボリュームＰＤ１２に対するキャッシュ属性を異ならせても良い。こうすることにより、例えば、仮想ボリュームＶＤａからの読み出し時にはキャッシュを許可し、仮想ボリュームＶＤｂからの読み出し時にはキャッシュを禁止するというように、読み出すクライアントに応じてキャッシュの可否を切り換えることが可能となる。

【００４６】

書き込みについても、同様の形式でキャッシュの可否が管理されている。キャッシュ可否については、計算機システムの動作時に動的に設定を変更してもよい。例えば、仮想ボリュームＶＤａを使用するクライアントの動作異常を検知した時、ノード装置２００は仮想ボリュームＶＤａに対応する全物理ボリュームについて、一時的に「キャッシュ不可」の属性を設定するようにしてもよい（図中のデータ領域Ｄａ参照）。

【００４７】

Ａ４．データ読出処理；

図７はデータ読み出し処理のフローチャートである。リモートブート時を例にとり、左側にクライアント１００の処理、中央にノード装置２００の処理、右側にストレージ装置３００の処理を示した。

【００４８】

ユーザが計算機の電源をオンにすると、ＩＤ入力モジュール１３１によって、ユーザＩＤ、パスワード等の入力画面が提示される。ユーザが、ユーザＩＤ、パスワードを入力し、ログインすると、クライアント１００は、ＩＰアドレスとともにこれらの情報を、ノード装置２００およびストレージ装置３００に送信する（ステップＳ１０）。ストレージ装置３００は、この情報に基づいてユーザ管理テーブル３３４を参照し、ユーザの認証を行う（ステップＳ３０）。正規なユーザであることが確認されると、ストレージ装置３００は、ユーザ管理テーブル３３４に、クライアント１００のＩＰアドレスを記録するとともに、ユーザに対応した仮想ボリューム名を通知する。この通知は、ノード装置２００を介してクライアント１００に送信される。以後、クライアント１００からのデータの読み出し要求等は、この仮想ボリューム名に基づいて行われる。

【００４９】

クライアント１００は、ＲＯＭに記録されているブートプログラム１３０を起動する（ステップＳ１１）。ブートプログラム１３０は、ブートに必要な初期化処理を行った上で、オペレーティングシステムのデータ読み出し要求をノード装置に対して送信する（ステップＳ１２）。

【００５０】

ノード装置２００は、読み出し要求を受信し、キャッシュ管理テーブル２３２を参照して、要求されたデータが、既にキャッシュ済みであるか否かを確認する（ステップＳ２０）。キャッシュ済みの場合は（ステップＳ２０）、キャッシュブロックから該当するデータを読み出し（ステップＳ２６）、クライアント１００に返信する（ステップＳ２５）。クライアント１００は、このデータを受信して起動処理を継続する（ステップＳ１３）。

【００５１】

要求されたデータがキャッシュされていない（以下、「未キャッシュ」と呼ぶ）場合（ステップＳ２０）、ノード装置２００はストレージ装置３００にデータの読み出しを要求する（ステップＳ２１）。ストレージ装置３００は、この要求に応じて、物理ブロック管理

テーブル３３２を参照して、物理ボリュームからデータを読み出し、ノード装置２００に返信する（ステップＳ３２）。これと併せて、ストレージ装置３００は、キャッシュ属性管理テーブル３３３を参照し、要求されたデータのキャッシュ可否の属性情報をノード装置２００に通知する（ステップＳ３２）。この属性情報は、要求されたデータのヘッダ等に埋め込んだ形で通知してもよいし、要求されたデータと分けて通知してもよい。

【００５２】

ノード装置２００は、ストレージ装置３００からデータおよび属性情報を受信すると、属性情報に基づいてキャッシュの可否を判断する（ステップＳ２２）。キャッシュ可と判断される場合には、受信したデータをキャッシュし（ステップＳ２３）、その格納先に応じてキャッシュ管理テーブル２３２を更新して（ステップＳ２４）、データをクライアントに返信する（ステップＳ２５）。キャッシュ不可と判断される場合には、上述のキャッシュ処理をスキップし、データをクライアントに返信する（ステップＳ２５）。クライアントはこうしてノード装置２００から送信されたデータを受信する（ステップＳ１３）。

【００５３】

以上の処理を繰り返し行うことにより、クライアント１００は、ブートに必要なファイルを取得し、リモートブートを完了することができる。図７ではリモートブート時の読み出し処理を示したが、リモートブートと無関係に通常のデータ読み出し時にも、データ読み出し要求（ステップＳ１２）からデータ受信（ステップＳ１３）の処理が同様にして行われる。

【００５４】

上述の実施例では、データの要求時にノード装置２００からストレージ装置３００にキャッシュの可否を問い合わせる場合を例示した。これに対し、ノード装置２００は、ストレージ装置３００から、予めキャッシュ属性管理テーブル３３３の内容をまとめて取得し、キャッシュ属性管理部２３４で管理するようにしてもよい。こうすることにより、データ要求のたびに、属性情報の問い合わせを行う必要性をなくすことができる。

【００５５】

Ａ５．データ書込処理；

図８はデータ書込処理のフローチャートである。左側にクライアント１００の処理、中央にノード装置２００の処理、右側にストレージ装置３００の処理を示した。クライアント１００は既に起動し、ノード装置２００、ストレージ装置３００とデータ授受が可能な状態で、クライアント１００の動作にソフトウェア上の異常が生じた場合のコアダンプの書き込みを例示した。

【００５６】

クライアント１００の異常制御モジュール１４１は、ソフトウェア上の動作異常を検出すると（ステップＳ４０）、異常の発生をＩＰアドレスとともにノード装置２００に通知する（ステップＳ４０）。ノード装置２００は、この情報に基づいてキャッシュ属性管理部２３４における属性情報を更新する（ステップＳ５０）。先に図６で示したように、このクライアントからの書き込みについては、全てキャッシュ不可という属性に変更するのである。

【００５７】

クライアント１００の異常制御モジュール１４１は、仮想ブロックを指定して、コアダンプを書き込むための要求をノード装置に送信する（ステップＳ４２）。ノード装置２００は、書き込み要求を受信し、キャッシュ属性管理部２３４を参照して、キャッシュの可否を判断する（ステップＳ５２）。クライアント１００の動作に異常が生じた場合には、ノード装置２００は、ステップＳ５０で更新された属性情報に基づき、キャッシュ不可と判断することになる。一方、クライアント１００に異常が生じる前、異常から回復した後、および別のクライアントからの書き込み要求時などには、指定された仮想ブロックの属性情報に基づき、ノード装置２００はキャッシュ可と判断する場合もある。

【００５８】

キャッシュ可と判断した場合（ステップＳ５２）、ノード装置は、クライアントから受信

したデータをキャッシュし（ステップＳ５３）、その格納先に応じてキャッシュ管理テーブル２３２を更新するとともに（ステップＳ５４）、データをストレージ装置３００に送信する（ステップＳ５５）。キャッシュ不可と判断した場合（ステップＳ５２）、ノード装置２００は、クライアントから受信したデータをキャッシュすることなく、ストレージ装置３００に送信する（ステップＳ５５）。ストレージ装置は、ノード装置２００からデータを受信して（ステップＳ６０）、指定された仮想ブロックに書き込みを行う。

【００５９】

以上の処理を繰り返し行うことにより、クライアント１００は、ストレージ装置３００の仮想ブロックにデータの書き込みを行うことができる。この際、異常が生じた場合のコアダンプについては、ノード装置へのキャッシュを回避することができる。

【００６０】

上述の実施例では、クライアント１００の異常通知により、ノード装置２００がキャッシュ可否の属性情報を更新する場合を例示した。これに対し、クライアント１００からノード装置２００に送信されるデータごとにキャッシュ可否の属性情報を添付するようにしてもよい。例えば、異常制御モジュール１４１は、コアダンプとして書き込むデータについては、全てキャッシュ不可という属性情報を添付し、ノード装置２００は、この属性情報に基づいてキャッシュの可否を判断するようにしてもよい。

【００６１】

Ａ６．変形例；

実施例では、ストレージ装置３００からのデータ読み込み時、およびストレージ装置３００への書き込み処理時にノード装置２００が一切、キャッシュを行わない例を示した。ノード装置２００は、キャッシュ可否に関わらず、揮発性メモリ２０２へのキャッシュを行い、メモリ２０２からディスク装置２１０へのデータ移転時にキャッシュの可否を判定するようにしてもよい。かかる制御は、ディスク装置２１０への書き込み時に、キャッシュ属性管理部２３４とキャッシュ管理テーブル２３２を参照することで実現することができる。つまり、キャッシュ管理テーブル２３２を参照することにより、メモリ２０２にキャッシュされているデータについて、仮想ボリュームおよびブロック番号を特定することができる。この情報に基づいて、キャッシュ属性管理部２３４を参照すれば、ディスク装置２１０へのキャッシュ可否を判定することができる。

【００６２】

変形例の制御を実現するための別の方法として、キャッシュ管理テーブル２３２に、図５の破線で示した「ディスク書込」という情報を設けても良い。「ディスク書込」は、ディスク装置２１０へのキャッシュ可否を表す属性情報であり、メモリ２０２へのキャッシュ時に、キャッシュ属性管理部２３４を参照することで設定することができる。このようにキャッシュ管理テーブル２３２において、格納されているデータごとに、ディスク書込を保存しておくことにより、ディスク装置２１０への移転時にキャッシュ属性管理部２３４を参照する必要がなくなり、キャッシュ可否を簡易に判定することが可能となる。

【００６３】

以上で説明した第１実施例によれば、データの読みだし、および書き込み時に、ノード装置２００におけるディスク装置２１０へのキャッシュの可否を切り替えることができる。コアダンプなど、機密情報を含むデータについては、ディスク装置２１０へのキャッシュを回避することができるため、ノード装置２００の盗難による情報漏洩を抑止することができる。ノード装置２００においてメモリ２０２へのキャッシュ自体も禁止する場合には、ノード装置２００への不正アクセスによる情報漏洩を抑止することも可能である。

【００６４】

Ｂ．第２実施例；

第１実施例では、属性情報に基づいて、ノード装置２００におけるキャッシュ可否を切り替える例を示した。第２実施例では、ノード装置が、ディスク装置へのキャッシュ時に、データを暗号化することでセキュリティを確保する例を示す。ノード装置への不正アクセスに対するセキュリティ確保という観点からは、メモリ２０２へのキャッシュ時にもデー

タを暗号化してもよい。ただし、本実施例においては、暗号化に要する処理負荷を軽減するため、メモリ２０２へのキャッシュ時には暗号化を行わず、ディスク装置２１０へのデータ移転時に暗号化する例を示す。

【００６５】

Ｂ１．機能ブロック；

図９は第２実施例としての計算機システムの機能ブロックを示す説明図である。第１実施例と同じ機能ブロックについては、同一の符号で示した。第２実施例では、キャッシュ属性管理部２３４、キャッシュ属性管理テーブル３３３に確保されている属性情報は、キャッシュの可否ではなく、暗号化の要否を示す情報となる。第２実施例では、ノード装置２００Ａに暗号化部２３５が新たに追加され、ストレージ装置３００Ａに鍵管理モジュール３３５が新たに追加される。暗号化部２３５は、ノード装置２００Ａにおいて、ディスク装置２１０へのキャッシュ時に、データの暗号化を行う機能、およびディスク装置２１０から読み出されたデータを復号する機能を奏する。鍵管理モジュール３３５は、ノード装置２００Ａに対して、暗号化に使用する鍵情報を管理・提供する機能を奏する。

【００６６】

Ｂ２．データ読出処理；

図１０は第２実施例におけるデータ読出処理のフローチャートである。左側にクライアント１００の処理、中央にノード装置２００Ａの処理、右側にストレージ装置３００Ａの処理を示した。第２実施例では、ノード装置２００Ａがネットワークに接続され、起動すると、ストレージ装置３００Ａは、ノード装置２００に対して鍵情報の送信を行う（ステップＳ３１）。ノード装置２００Ａは、この鍵情報を暗号化部２３５に記憶する（ステップＳ７０）。

【００６７】

本実施例においては、鍵情報は、ノード装置ごとに固有の情報として設定される。ストレージ装置３００は、ネットワークに接続されている複数のノード装置について、提供した鍵情報を一元的に管理し、新たに接続されたノード装置に対しては、これらのいずれとも異なる鍵情報を生成する。鍵情報は、例えば、シリアル番号などの規則的なものであってもよいし、乱数を用いた不規則な情報であってもよい。各ノード装置に一意な鍵情報を提供することにより、いわゆる悪意の第三者がノード装置になりすまして鍵情報を取得した場合でも、各ノード装置にキャッシュされたデータのセキュリティを確保することができる利点がある。

【００６８】

クライアント１００がノード装置２００Ａに対して、データの読みだし要求を送信すると（ステップＳ１２）、ノード装置２００Ａは、要求されたデータが、キャッシュ済みであるか否かを判断する（ステップＳ７１）。キャッシュ済みである場合には、メモリ２０２またはディスク装置２１０からデータを読みだし（ステップＳ７２）、暗号化されている場合には、復号して（ステップＳ７３、Ｓ７４）、クライアント１００に返信する（ステップＳ８１、ステップＳ１３）。

【００６９】

要求されたデータがキャッシュされていない場合（ステップＳ７１）、ノード装置２００Ａは、ストレージ装置３００Ａにデータを要求する（ステップＳ７５）。ストレージ装置３００Ａは、この要求に応じて、データをノード装置２００Ａに返信する（ステップＳ３２）。これと併せて、ストレージ装置３００は、要求されたデータについて暗号化の要否を表す属性情報をノード装置２００Ａに通知する（ステップＳ３２）。この属性情報は、要求されたデータのヘッダ等に埋め込んだ形で通知してもよいし、要求されたデータと分けて通知してもよい。

【００７０】

ノード装置２００Ａは、ストレージ装置３００Ａからデータおよび属性情報を受信すると、データのキャッシュ先を判断する（ステップＳ７６）。メモリ２０２にキャッシュ可能な領域があいている場合には、メモリ２０２へのキャッシュを行う。キャッシュ可能な領

域が空いていない場合には、ディスク装置２１０へのキャッシュが必要と判断する。この判断は、ストレージ装置３００Ａからのデータ受信時に行っても良いし、いわゆるＬＲＵ（Ｌｅａｓｔ　Ｒｅｃｅｎｔｌｙ　Ｕｓｅｄ）に基づいて処理してもよい。ＬＲＵとは、メモリ２０２にキャッシュされているデータのうち、アクセス頻度の低いデータを優先的にディスク装置２１０に移転する方法である。ＬＲＵで処理する場合には、ストレージ装置３００Ａから受信したデータは、無条件にメモリ２０２にキャッシュされ、代わりに、従前にメモリ２０２に蓄積されていたデータのいずれかがディスク装置２１０に移転されることになる。

【００７１】

ノード装置２００Ａは、ディスク装置２１０にキャッシュすべきと判断されたデータについては（ステップＳ７６）、属性情報に基づいて暗号化の要否を判断する（ステップＳ７７）。暗号化が要求されている場合には、暗号化を行い（ステップＳ７８）、キャッシュする（ステップＳ７９）。暗号化が要求されていないデータ、またはメモリ２０２にキャッシュすべきデータについては（ステップＳ７６）、暗号化を行わずにキャッシュする。ノード装置２００Ａは、これらのキャッシュの格納先に応じて、キャッシュ管理テーブル２３２を更新し（ステップＳ８０）、データをクライアント１００に返信する（ステップＳ２５）。クライアントはこうしてノード装置２００から送信されたデータを受信する（ステップＳ１３）。

【００７２】

以上の処理を繰り返し行うことにより、クライアント１００は、必要なファイルをストレージ装置３００Ａから読み出すことができる。この際、機密情報を含むデータについては、属性情報の設定により、ディスク装置２１０へのキャッシュ時に暗号化することができるため、セキュリティを確保することができる。第２実施例においても第１実施例と同様、ノード装置２００Ａが、暗号要否の属性情報をストレージ装置３００Ａから、予めまとめて取得し、管理するようにしてもよい。

【００７３】

Ｂ３．データ書込処理；

図１１は第２実施例におけるデータ書込処理のフローチャートである。左側にクライアント１００の処理、中央にノード装置２００Ａの処理、右側にストレージ装置３００Ａの処理を、コアダンプの書き込みを例にとって示した。

【００７４】

読みだし処理と同じく、ノード装置２００Ａは起動時に、ストレージ装置から鍵情報を受信し、記憶している（ステップＳ６０、Ｓ３１）。また、第１実施例と同様、クライアント１００の異常制御モジュール１４１は、ソフトウェア上の動作異常を検出すると（ステップＳ４０）、異常の発生をＩＰアドレスとともにノード装置２００に通知する（ステップＳ４０）。ノード装置２００は、この情報に基づいてキャッシュ属性管理部２３４における属性情報を暗号化要という内容に更新する（ステップＳ５０）。

【００７５】

この状態で、クライアント１００がコアダンプを書き込むための要求をノード装置２００Ａに送信すると（ステップＳ４２）、ノード装置２００Ａは、読みだし処理と同様、受信したデータのキャッシュ先を判断する。ノード装置２００Ａは、ディスク装置２１０にキャッシュすべきと判断されたデータについて、属性情報で指定された暗号化の要否に基づいて暗号化を行った上で、キャッシュする（ステップＳ９０～Ｓ９３）。暗号化が要求されていないデータ、またはメモリ２０２にキャッシュすべきデータについては（ステップＳ９０）、暗号化を行わずにキャッシュする（ステップＳ９４）。ノード装置２００Ａは、これらのキャッシュの格納先に応じて、キャッシュ管理テーブル２３２を更新し（ステップＳ９４）、データをストレージ装置３００Ａに送信する（ステップＳ９５）。ストレージ装置３００Ａは、ノード装置２００Ａからデータを受信して（ステップＳ６０）、指定された仮想ブロックに書き込みを行う。以上の処理を繰り返し行うことにより、クライアント１００は、ストレージ装置３００Ａの仮想ブロックにデータの書き込みを行うこと

ができる。

【００７６】

第２実施例によれば、ノード装置２００Ａにおいて、ディスク装置への書き込み時に暗号化することにより、盗難によるデータの漏洩を抑止することができる。メモリ２０２への書き込み時には暗号化を行わないため、頻繁に使用されるデータについては、メモリ２０２にキャッシュしておくことにより、読みだし、書き込みの速度を確保することもできる。

【００７７】

第２実施例では、鍵情報をストレージ装置３００Ａから提供したが、鍵情報は厳密に一意である必要はなく、ノード装置２００Ａで生成しても構わない。第２実施例では、属性情報によって暗号化の要否を制御する場合を例示したが、ディスク装置２１０へのキャッシュ時には必ず暗号化するようにしてもよい。

【００７８】

Ｃ．第３実施例；

第３実施例では、キャッシュ可とされているデータについては、クライアントとストレージ装置間でノード装置を介してデータの授受を行い、キャッシュ不可とされているデータについては、両者間で直接データの授受を行う場合を例示する。

【００７９】

Ｃ１．機能ブロック；

図１２は第３実施例としての計算機システムの機能ブロックを示す説明図である。第１実施例と同じ機能ブロックについては、同一の符号で示した。第３実施例では、クライアント１００に提供されるオペレーティングシステム１４０Ｂの内容が、第１実施例と相違する。オペレーティングシステム１４０Ｂは、ＳＣＳＩコマンドの送出先としてノード装置２００Ｂのアドレス（以下、「ノードアドレス」と称する）１４２を保持している。オペレーティングシステム１４０Ｂには、動作異常時にコアダンプの書き込みなどを行う異常制御モジュール１４１Ｂが供えられており、この異常制御モジュール１４１Ｂは、コアダンプの書き込み先としてストレージ装置３００Ｂのアドレス（以下、「ストレージアドレス」と称する）１４３を保持している。これらの機能ブロックにより、クライアントは、異常時と正常時でデータの書き込み先を切り替えることができる。

【００８０】

ストレージ装置３００Ｂにおいては、アクセス制御モジュール３３０Ｂが、キャッシュ属性に基づいてデータの送信先を切り替える。キャッシュ可とされているデータについてはノード装置に送信し、キャッシュ不可とされているデータはクライアント１００に直接送信する。第３実施例では、キャッシュ不可とされているデータについては、クライアント１００とストレージ装置３００Ｂの間で直接送受信されるから、ノード装置２００Ｂにおいて、キャッシュ制御モジュール２３０Ｂは、クライアント１００およびストレージ装置３００Ｂから受信したデータについては、キャッシュ可否を考慮することなく、データをキャッシュする。

【００８１】

Ｃ２．データ読出処理；

図１３は第３実施例におけるデータ読み出し処理のフローチャートである。リモートブート時を例にとり、左側にクライアント１００の処理、中央にノード装置２００Ｂの処理、右側にストレージ装置３００Ｂの処理を示した。

【００８２】

ユーザが計算機の電源をオンにすると、第１実施例と同様の手順でログインおよびブートプログラムの起動が行われる（ステップＳ１０、Ｓ３０、Ｓ１１）。クライアント１００は、仮想ボリュームを指定して、データの読み出し要求をノード装置２００Ｂに対して送信する（ステップＳ１２）。

【００８３】

ノード装置２００Ｂは、要求されたデータがキャッシュ済みの場合は（ステップＳ１００

)、キャッシュブロックから該当するデータを読み出し（ステップＳ１０１）、クライアント１００に返信する（ステップＳ１０５）。クライアント１００は、このデータを受信して起動処理を継続する（ステップＳ１３）。

【００８４】

要求されたデータが未キャッシュの場合（ステップＳ１００）、ノード装置２００Ｂはストレージ装置３００Ｂにデータの読み出しを要求する（ステップＳ１０２）。ストレージ装置３００Ｂは、キャッシュ属性管理テーブル３３３において、要求されたデータがキャッシュ可と設定されている場合には（ステップＳ１１０）、データをノード装置に返信する（ステップＳ１１１）。ノード装置２００Ｂは、このデータをキャッシュし（ステップＳ１０３）、キャッシュ管理テーブル２３２を更新して（ステップＳ２４）、データをクライアントに返信する（ステップＳ１０５）。

【００８５】

要求されたデータがキャッシュ不可の場合には（ステップＳ１１０）、ストレージ装置３００Ｂは、データをクライアントに直接、返信する（ステップＳ１１２）。クライアントはこうしてノード装置２００から送信されたデータを受信する（ステップＳ１３）。ストレージ装置３００Ｂは、ノード装置２００Ｂに対して、クライアントへの返信を行った旨の通知を行っても良い。ノード装置２００Ｂが、データの要求（ステップＳ１０２）へのレスポンスを待っている場合には、この通知により、レスポンス待ちの状態から解放することができる。

【００８６】

上記実施例において、ストレージ装置３００Ｂからクライアントに直接、データを送信する場合、クライアントのアドレスは種々の方法で特定することができる。例えば、ストレージ装置３００Ｂは、最初のログイン時にユーザ管理テーブル３３４に記録されているＩＰアドレスを利用してもよい。別の例として、データの要求時に、ノード装置２００Ｂからストレージ装置３００Ｂに対して、クライアント１００のアドレス情報を通知するようにしてもよい。

【００８７】

Ｃ３．データ書込処理；

図１４はデータ書込処理のフローチャートである。左側にクライアント１００の処理、中央にノード装置２００Ｂの処理、右側にストレージ装置３００Ｂの処理を、コアダンプの書き込みを例にとって示した。クライアント１００は、異常が生じていない場合には（ステップＳ１２０）、ノード装置２００に対してデータの書き込み要求を送信する（ステップＳ１２１）。ノード装置２００Ｂは、この要求を受信すると、データをキャッシュし（ステップＳ１３）、キャッシュ管理テーブル２３２を更新する（ステップＳ１３１）。その後、キャッシュされている書き込みデータが所定量に達したなど、予め設定されたタイミングで、データをストレージ装置３００Ｂに送信する（ステップＳ１３２）。ストレージ装置３００Ｂは、このデータを受信して、指定された仮想ブロックにデータの書き込みを行う（ステップＳ１４０）。

【００８８】

一方、異常制御モジュール１４１が、ソフトウェア上の動作異常を検出すると（ステップＳ１２０）、データの送信先をノード装置２００のアドレスからストレージ装置に切り替えることにより、ストレージ装置３００Ｂに対して直接、データの書き込み要求を送信する（ステップＳ１２２）。ストレージ装置３００Ｂは、このデータを受信して、指定された仮想ブロックにデータの書き込みを行う（ステップＳ１４０）。

【００８９】

第３実施例によれば、キャッシュの可否によって、ノード装置２００Ｂを迂回してデータを送受信することができる。従って、ノード装置の盗難や不正アクセスに対し、データのセキュリティを確保することができる。

【００９０】

以上、本発明の種々の実施例について説明したが、本発明はこれらの実施例に限定されず

、その趣旨を逸脱しない範囲で種々の構成を採ることができることはいうまでもない。

【００９１】

【発明の効果】

本発明によれば、ノード装置へのキャッシュ自体の回避、キャッシュ時の暗号化、ノード装置を迂回したデータ授受などの方法により、機密情報を含んだデータ等がノード装置のディスク装置に、原状でキャッシュされることを回避することができる。従って、ノード装置の盗難、不正アクセスに対して、これらのデータのセキュリティを確保することができる。

【図面の簡単な説明】

【図１】第１実施例としての計算機システムの構成を示す説明図である。

【図２】計算機システムを構成する各装置内の機能ブロックを示す説明図である。

【図３】ユーザ管理テーブル３３４の構造を示す説明図である。

【図４】物理ブロック管理テーブル３３２の構造を示す説明図である。

【図５】キャッシュ管理テーブル２３２の構造を示す説明図である。

【図６】キャッシュ属性管理テーブル３３３の構造を示す説明図である。

【図７】データ読み出し処理のフローチャートである。

【図８】データ書込処理のフローチャートである。

【図９】第２実施例としての計算機システムの機能ブロックを示す説明図である。

【図１０】第２実施例におけるデータ読出処理のフローチャートである。

【図１１】第２実施例におけるデータ書込処理のフローチャートである。

【図１２】第３実施例としての計算機システムの機能ブロックを示す説明図である。

【図１３】第３実施例におけるデータ読み出し処理のフローチャートである。

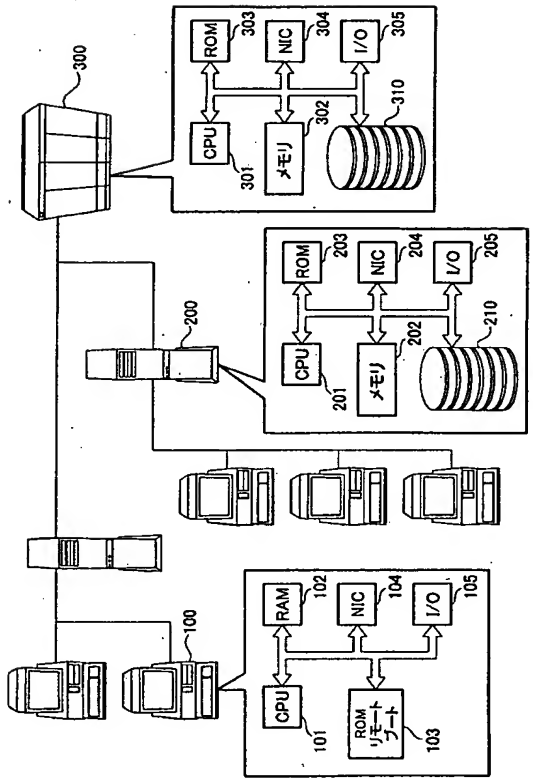
【図１４】データ書込処理のフローチャートである。

【符号の説明】

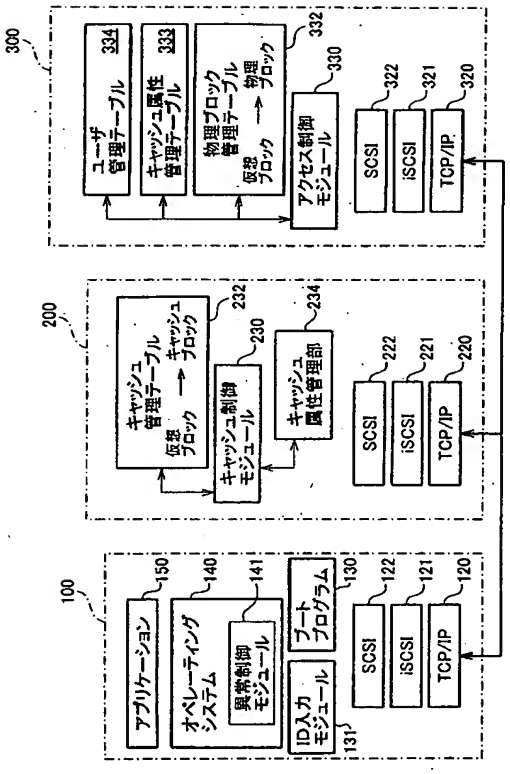
１００…クライアント
１０１、２０１、３０１…ＣＰＵ
１０２、２０２、３０２…メモリ（ＲＡＭ）
１０３、２０３、３０３…ＲＯＭ
１０４、２０４、３０４…ＮＩＣ
１０５、２０５、３０５…Ｉ／Ｏ
１２０、２２０、３２０…ＴＣＰ／ＩＰモジュール
１２１、２２１、３２１…ｉＳＣＳＩモジュール
１２２、２２２、３２２…ＳＣＳＩモジュール
１３０…ブートプログラム
１３１…ＩＤ入力モジュール
１４０、１４０Ｂ…オペレーティングシステム
１４１、１４１Ｂ…異常制御モジュール
１４２…ノードアドレス
１４３…ストレージアドレス
１５０…アプリケーション
２００、２００Ａ、２００Ｂ…ノード装置
２１０、３１０…ディスク装置
２３０、２３０Ｂ…キャッシュ制御モジュール
２３２…キャッシュ管理テーブル
２３４…キャッシュ属性管理部
２３５…暗号化部
３００、３００Ａ、３００Ｂ…ストレージ装置
３３０、３３０Ｂ…アクセス制御モジュール
３３２…物理ブロック管理テーブル
３３３…キャッシュ属性管理テーブル

３３４…ユーザ管理テーブル

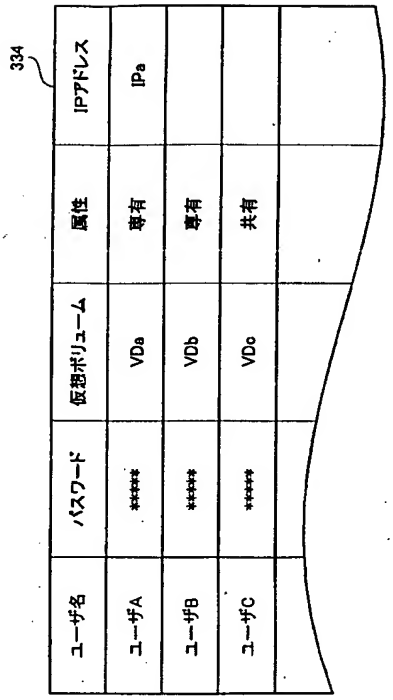
【図１】



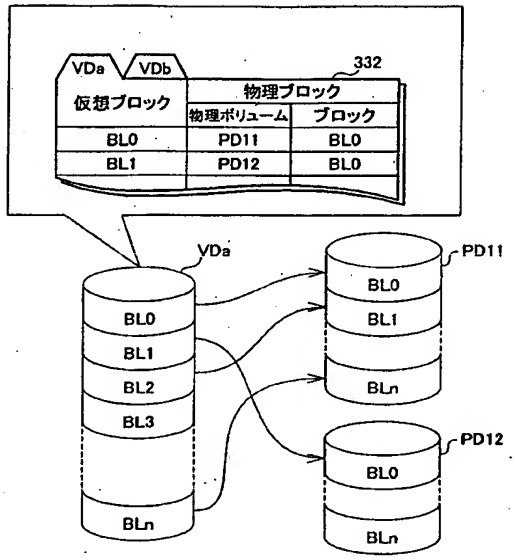
【図２】



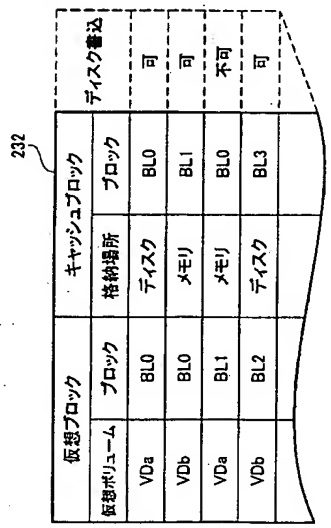
【図３】


334
ユーザ名
パスワード
仮想ボリューム
属性
IPアドレス
ユーザA
*****
VDa
専有
IPa
ユーザB
*****
VDb
専有
ユーザC
*****
VDc
共有

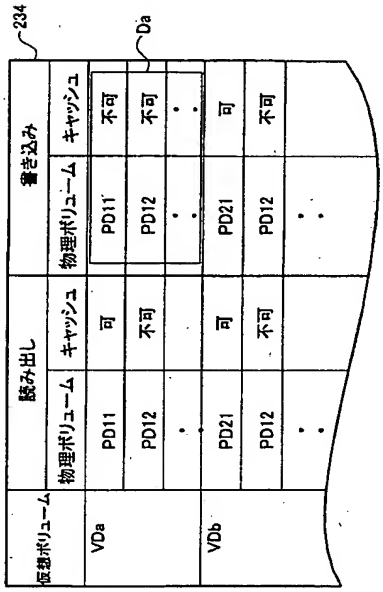
【図４】


VDa
VDb
仮想ブロック
物理ブロック
物理ボリューム
ブロック
332
BL0
PD11
BL0
BL1
PD12
BL0
VDa
BL0
BL1
BL2
BL3
BLn
PD11
BL0
BL1
BLn
PD12
BL0
BLn

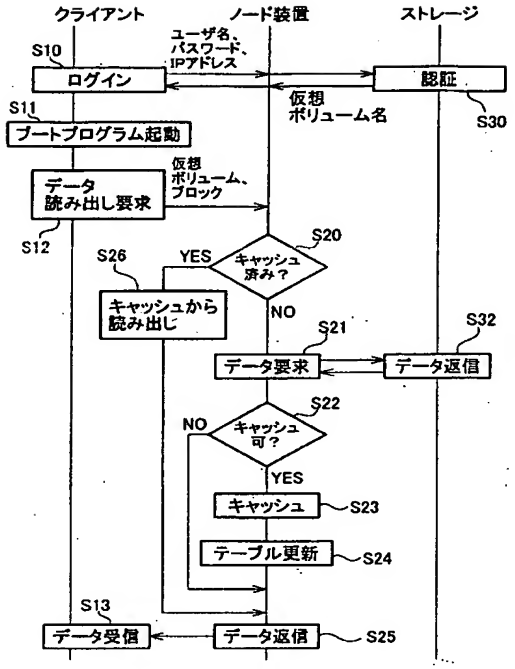
【図５】


232
仮想ブロック
仮想ボリューム
ブロック
キャッシュブロック
格納場所
ブロック
ディスク書込
VDa
BL0
ディスク
BL0
可
VDb
BL0
メモリ
BL1
可
VDa
BL1
メモリ
BL0
不可
VDb
BL2
ディスク
BL3
可

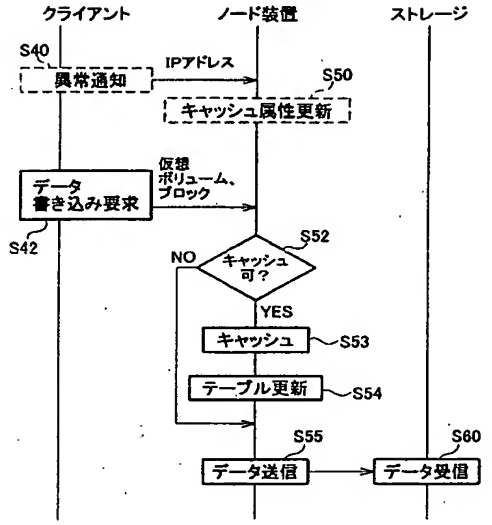
【図６】


234
Da
仮想ボリューム
読み出し
物理ボリューム
キャッシュ
書き込み
物理ボリューム
キャッシュ
VDa
PD11
可
PD11
不可
PD12
不可
PD12
PD12
VDb
PD21
可
PD21
可
PD12
不可
PD12
不可

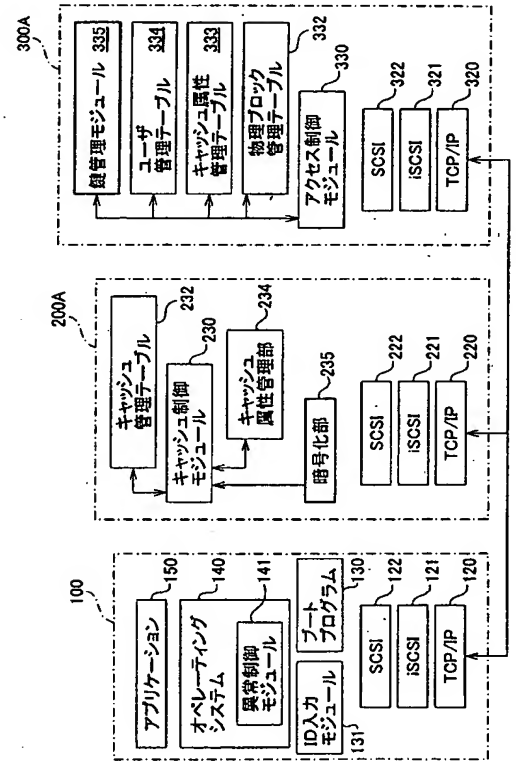
【図７】


クライアント
ノード装置
ストレージ
S10
ユーザ名、パスワード、IPアドレス
ログイン
認証
仮想ボリューム名
S30
S11
ブートプログラム起動
データ読み出し要求
仮想ボリューム、ブロック
S12
S20
キャッシュ済み？
YES
S26
キャッシュから読み出し
NO
S21
S32
データ要求
データ返信
S22
キャッシュ可？
NO
YES
キャッシュ
S23
テーブル更新
S24
S13
データ受信
データ返信
S25

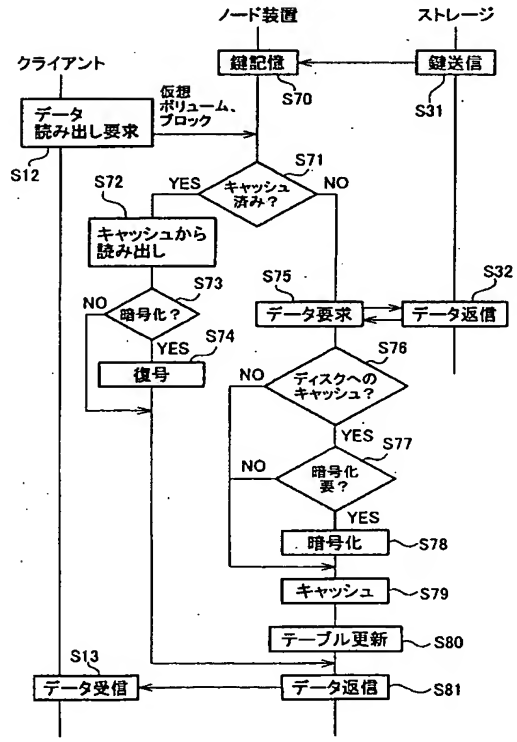
【図８】


クライアント
ノード装置
ストレージ
S40
異常通知
IPアドレス
S50
キャッシュ属性更新
データ書き込み要求
仮想ボリューム、ブロック
S42
S52
NO
キャッシュ可？
YES
キャッシュ
S53
テーブル更新
S54
S55
S60
データ送信
データ受信

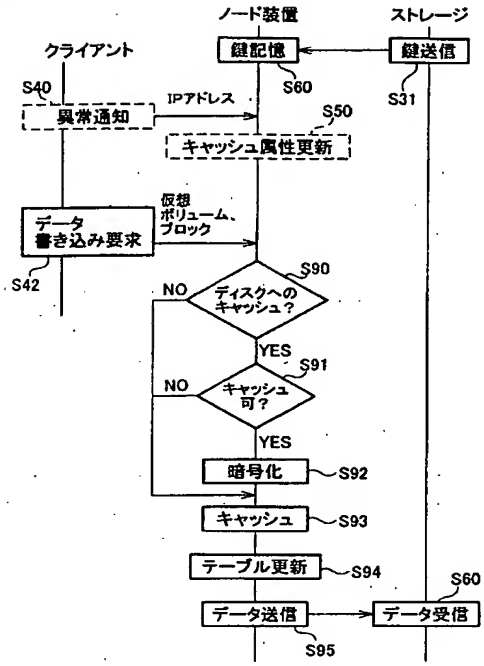
【図９】


300A
鍵管理モジュール 335
ユーザ管理テーブル 334
キャッシュ属性管理テーブル 333
物理ブロック管理テーブル 332
アクセス制御モジュール 330
SCSI 322
iSCSI 321
TCP/IP 320
200A
キャッシュ管理テーブル 232
キャッシュ制御モジュール 230
キャッシュ属性管理部 234
暗号化部 235
SCSI 222
iSCSI 221
TCP/IP 220
100
アプリケーション 150
オペレーティングシステム 140
異常制御モジュール 141
ブートプログラム 130
ID入力モジュール 131
SCSI 122
iSCSI 121
TCP/IP 120

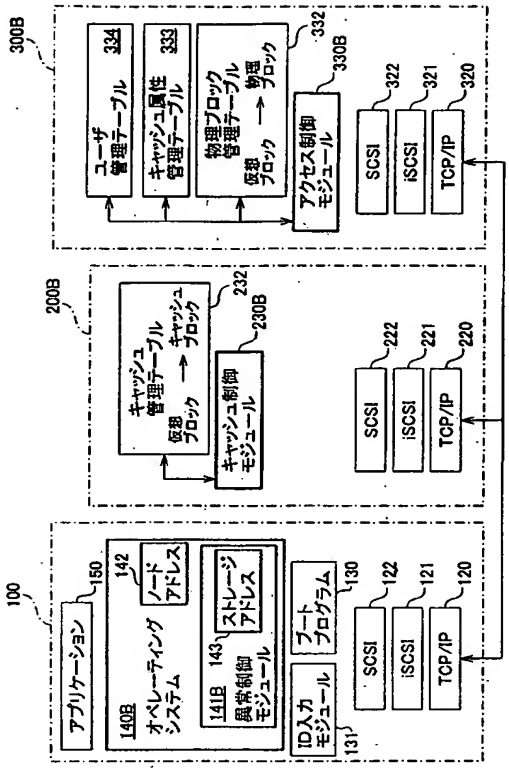
【図１０】


ノード装置
ストレージ
クライアント
鍵記憶
鍵送信
S70
S31
データ読み出し要求
仮想ボリューム、ブロック
S12
S71
S72
YES
キャッシュ済み？
NO
キャッシュから読み出し
S73
NO
暗号化？
YES
S74
復号
S75
S32
データ要求
データ返信
S76
NO
ディスクへのキャッシュ？
YES
S77
NO
暗号化要？
YES
暗号化
S78
キャッシュ
S79
テーブル更新
S80
S13
データ受信
データ返信
S81

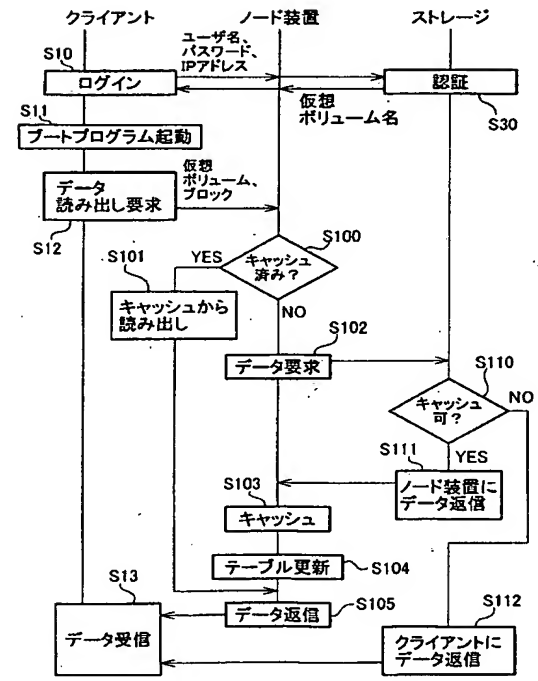
【図１１】



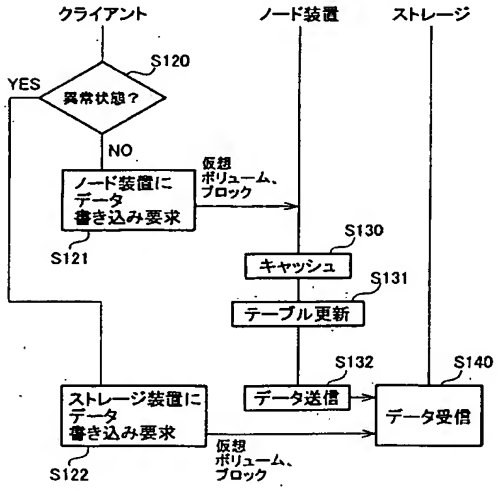
【図１２】



【図１３】



【図１４】



フロントページの続き

(72)発明者　橋本　尚
神奈川県横浜市戸塚区戸塚町５０３０番地　株式会社日立製作所ソフトウェア事業部内

Ｆターム(参考) 5B082 GA11 HA02